

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/031,496-D
Source: TPW/C
Date Processed by STIC: 12-2-04

ENTERED



IFW16

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/031,496D

DATE: 12/02/2004
TIME: 14:41:47

Input Set : A:\seq 4 99-45.ST25.txt
Output Set: N:\CRF4\12022004\J031496D.raw

```

3 <110> APPLICANT: National Renewable Energy Laboratory
5 <120> TITLE OF INVENTION: Cellobiohydrolase I Gene and Improved Variants
7 <130> FILE REFERENCE: NREL 99-45
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/031,496D
C--> 9 <141> CURRENT FILING DATE: 2002-01-14
9 <160> NUMBER OF SEQ ID NOS: 96
11 <170> SOFTWARE: PatentIn version 3.3
13 <210> SEQ ID NO: 1
14 <211> LENGTH: 28
15 <212> TYPE: DNA
16 <213> ORGANISM: Artificial
18 <220> FEATURE:
19 <223> OTHER INFORMATION: Primer for PCR
21 <400> SEQUENCE: 1
22 agagagtcta gacacggaggc ttacagggc 28
25 <210> SEQ ID NO: 2
26 <211> LENGTH: 35
27 <212> TYPE: DNA
28 <213> ORGANISM: Artificial
30 <220> FEATURE:
31 <223> OTHER INFORMATION: Primer for PCR
33 <400> SEQUENCE: 2
34 aaagaagcgc ggccgcgcct gcactctcca atcggg 35
37 <210> SEQ ID NO: 3
38 <211> LENGTH: 24
39 <212> TYPE: DNA
40 <213> ORGANISM: Artificial
42 <220> FEATURE:
43 <223> OTHER INFORMATION: Primer for PCR
45 <400> SEQUENCE: 3
46 ggcggaaacc cgcctggcac cacc 24
49 <210> SEQ ID NO: 4
50 <211> LENGTH: 1551
51 <212> TYPE: DNA
52 <213> ORGANISM: Trichoderma reesei
55 <220> FEATURE:
56 <221> NAME/KEY: CDS
57 <222> LOCATION: (1)..(1545)
59 <220> FEATURE:
60 <221> NAME/KEY: misc_feature
61 <222> LOCATION: (1)..(51)
63 <220> FEATURE:
64 <221> NAME/KEY: misc_feature

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/031,496D

DATE: 12/02/2004
TIME: 14:41:47

Input Set : A:\seq 4 99-45.ST25.txt
Output Set: N:\CRF4\12022004\J031496D.raw

65 <222> LOCATION: (52)..(1344)
 67 <220> FEATURE:
 68 <221> NAME/KEY: misc_binding
 69 <222> LOCATION: (1436)..(1550)
 71 <400> SEQUENCE: 4
 72 atg tat cgg aag ttg gcc gtc atc tcg gcc ttc ttg gcc aca gct cgt 48
 73 Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg
 74 1 5 10 15
 76 gct cag tcg gcc tgc act ctc caa tcg gag act cac ccg cct ctg aca 96
 77 Ala Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr
 78 20 25 30
 80 tgg cag aaa tgc tcg tct ggt ggc acg tgc act caa cag aca ggc tcc 144
 81 Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser
 82 35 40 45
 84 gtg gtc atc gac gcc aac tgg cgc tgg act cac gct acg aac agc agc 192
 85 Val Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser
 86 50 55 60
 88 acg aac tgc tac gat ggc aac act tgg agc tgc acc cta tgt cct gac 240
 89 Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp
 90 65 70 75 80
 92 aac gag acc tgc gcg aag aac tgc tgt ctg gac ggt gcc gcc tac gcg 288
 93 Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala
 94 85 90 95
 96 tcc acg tac gga gtt acc acg agc ggt aac agc ctc tcc att ggc ttt 336
 97 Ser Thr Tyr Gly Val Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe
 98 100 105 110
 100 gtc acc cag tct gcg cag aag aac gtt ggc gct cgc ctt tac ctt atg 384
 101 Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met
 102 115 120 125
 104 gcg agc gac acc tac cag gaa ttc acc ctg ctt ggc aac gag ttc 432
 105 Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe
 106 130 135 140
 108 tct ttc gat gtt gtt tcg cag ctg ccg tgc ggc ttg aac gga gct 480
 109 Ser Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala
 110 145 150 155 160
 112 ctc tac ttc gtg tcc atg gac gcg gat ggt ggc gtg agc aag tat ccc 528
 113 Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro
 114 165 170 175
 116 acc aac acc gct gcg gcc aag tac ggc acg ggg tac tgt gac agc cag 576
 117 Thr Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln
 118 180 185 190
 120 tgt ccc cgc gat ctg aag ttc atc aat ggc cag gcc aac gtt gag ggc 624
 121 Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly
 122 195 200 205
 124 tgg gag ccg tca tcc aac aac gcg aac acg ggc att gga gga cac gga 672
 125 Trp Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly
 126 210 215 220
 128 agc tgc tgc tct gag atg gat atc tgg gag gcc aac tcc atc tcc gag 720
 129 Ser Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/031,496D

DATE: 12/02/2004
TIME: 14:41:47

Input Set : A:\seq 4 99-45.ST25.txt
Output Set: N:\CRF4\12022004\J031496D.raw

130	225	230	235	240	
132	gct ctt acc ccc cac cct tgc acg act gtc ggc cag gag atc tgc gag				768
133	Ala Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu				
134	245	250	255		
136	ggt gat ggg tgc ggc gga act tac tcc gat aac aga tat ggc ggc act				816
137	Gly Asp Gly Cys Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr				
138	260	265	270		
140	tgc gat ccc gat ggc tgc gac tgg aac cca tac cgc ctg ggc aac acc				864
141	Cys Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr				
142	275	280	285		
144	agc ttc tac ggc cct ggc tca agc ttt acc ctc gat acc acc aag aaa				912
145	Ser Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys				
146	290	295	300		
148	ttg acc gtt gtc acc cag ttc gag acg tcg ggt gcc atc aac cga tac				960
149	Leu Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr				
150	305	310	315	320	
152	tat gtc cag aat ggc gtc act ttc cag cag ccc aac gcc gag ctt ggt				1008
153	Tyr Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly				
154	325	330	335		
156	agt tac tct ggc aac gag ctc aac gat gat tac tgc aca gct gag gag				1056
157	Ser Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu				
158	340	345	350		
160	gca gaa ttc ggc gga tcc tct ttc tca gac aag ggc ggc ctg act cag				1104
161	Ala Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Leu Thr Gln				
162	355	360	365		
164	ttc aag aag gct acc tct ggc ggc atg gtt ctg gtc atg agt ctg tgg				1152
165	Phe Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp				
166	370	375	380		
168	gat gat tac tac gcc aac atg ctg tgg gac tcc acc tac ccg aca				1200
169	Asp Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr				
170	385	390	395	400	
172	aac gag acc tcc tcc aca ccc ggt gcc gtg cgc gga agc tgc tcc acc				1248
173	Asn Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr				
174	405	410	415		
176	agc tcc ggt gtc cct gct cag gtc gaa tct cag tct ccc aac gcc aag				1296
177	Ser Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys				
178	420	425	430		
180	gtc acc ttc tcc aac atc aag ttc gga ccc att ggc agc acc ggc aac				1344
181	Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn				
182	435	440	445		
184	cct agc ggc ggc aac cct ccc ggc gga aac ccg cct ggc acc acc acc				1392
185	Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro Gly Thr Thr Thr				
186	450	455	460		
188	acc cgc cgc cca gcc act acc act gga agc tct ccc gga cct acc cag				1440
189	Thr Arg Arg Pro Ala Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln				
190	465	470	475	480	
192	tct cac tac ggc cag tgc ggc ggt att ggc tac agc ggc ccc acg gtc				1488
193	Ser His Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val				
194	485	490	495		

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,496D

DATE: 12/02/2004

TIME: 14:41:47

Input Set : A:\seq 4 99-45.ST25.txt

Output Set: N:\CRF4\12022004\J031496D.raw

196 tgc gcc agc ggc aca act tgc cag gtc ctg aac cct tac tac tct cag 1536
 197 Cys Ala Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln
 198 500 505 510
 200 tgc ctg taa agctcc 1551
 201 Cys Leu
 205 <210> SEQ ID NO: 5
 206 <211> LENGTH: 514
 207 <212> TYPE: PRT
 208 <213> ORGANISM: Trichoderma reesei
 210 <400> SEQUENCE: 5
 212 Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg
 213 1 5 10 15
 216 Ala Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr
 217 20 25 30
 220 Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser
 221 35 40 45
 224 Val Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser
 225 50 55 60
 228 Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp
 229 65 70 75 80
 232 Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala
 233 85 90 95
 236 Ser Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe
 237 100 105 110
 240 Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met
 241 115 120 125
 244 Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe
 245 130 135 140
 248 Ser Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala
 249 145 150 155 160
 252 Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro
 253 165 170 175
 256 Thr Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln
 257 180 185 190
 260 Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly
 261 195 200 205
 264 Trp Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly
 265 210 215 220
 268 Ser Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu
 269 225 230 235 240
 272 Ala Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu
 273 245 250 255
 276 Gly Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr
 277 260 265 270
 280 Cys Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr
 281 275 280 285
 284 Ser Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys
 285 290 295 300
 288 Leu Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,496D

DATE: 12/02/2004

TIME: 14:41:47

Input Set : A:\seq 4 99-45.ST25.txt

Output Set: N:\CRF4\12022004\J031496D.raw

289	305	310	315	320
292	Tyr Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly			
293		325	330	335
296	Ser Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu			
297		340	345	350
300	Ala Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln			
301		355	360	365
304	Phe Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp			
305		370	375	380
308	Asp Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr			
309		385	390	395
312	Asn Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr			
313		405	410	415
316	Ser Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys			
317		420	425	430
320	Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn			
321		435	440	445
324	Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro Gly Thr Thr Thr			
325		450	455	460
328	Thr Arg Arg Pro Ala Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln			
329		465	470	475
332	Ser His Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val			
333		485	490	495
336	Cys Ala Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln			
337		500	505	510
340	Cys Leu			
344	<210> SEQ ID NO: 6			
345	<211> LENGTH: 45			
346	<212> TYPE: DNA			
347	<213> ORGANISM: Artificial			
349	<220> FEATURE:			
350	<223> OTHER INFORMATION: Primer for PCR			
352	<400> SEQUENCE: 6			
353	cctccggcg gaaaccggcc tggcaccacc accacccgccc gccca			45
356	<210> SEQ ID NO: 7			
357	<211> LENGTH: 32			
358	<212> TYPE: DNA			
359	<213> ORGANISM: Artificial			
361	<220> FEATURE:			
362	<223> OTHER INFORMATION: Primer for PCR			
364	<400> SEQUENCE: 7			
365	ggactcacgc tacggccagc agcacgaact gc			32
368	<210> SEQ ID NO: 8			
369	<211> LENGTH: 36			
370	<212> TYPE: DNA			
371	<213> ORGANISM: Artificial			
373	<220> FEATURE:			
374	<223> OTHER INFORMATION: Primer for PCR			
376	<400> SEQUENCE: 8			

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 12/02/2004
PATENT APPLICATION: US/10/031,496D TIME: 14:41:48

Input Set : A:\seq 4 99-45.ST25.txt
Output Set: N:\CRF4\12022004\J031496D.raw

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:1,2,3,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29
Seq#:30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53
Seq#:54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77
Seq#:78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96

VERIFICATION SUMMARY DATE: 12/02/2004
PATENT APPLICATION: US/10/031,496D TIME: 14:41:48

Input Set : A:\seq 4 99-45.ST25.txt
Output Set: N:\CRF4\12022004\J031496D.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date